

In the Specification

Please replace paragraph 0005 with the following amended paragraph:

Figure 14[[1]] illustrates one way in which a microarray with tag-probes could be used to screen a protein library, with no cloning needed. To a protein-encoding mRNA a 5' tag sequence and a 3' ribosome-blocking sequence are attached (A). In a pool of such molecules, such as a randomly mutated gene library, each mRNA is paired with a unique tag and all have the same 3' sequence. Following in-vitro translation either on a microarray or in a test tube, the nascent protein remains attached to the mRNA (B), as in the technique of ribosome display (see, e.g., Hanes, et al. (2000) *Methods Enzymol* 328:404). During hybridization the tag directs each mRNA or mRNA-protein complex to a particular address on the Tag probe array (C), where all the proteins are screened simultaneously for activity (D). Appropriate detection methods identify proteins of interest (E), and the corresponding tag is known by the address on the array. Finally, the corresponding genes can be captured by RT-PCR of the mRNA pool, either from the mRNA on the array or from another aliquot, using a universal reverse primer and each identified Tag sequence as a forward primer. The genes can then be subjected to further screening or another round of mutagenesis.

Please insert the following paragraph as paragraph 0017 after paragraph 0016:

Figure 10. Gene expression monitoring with oligonucleotide arrays. A. An image of a hybridized 1.28 X 1.28 cm HuGeneFL array, with 20 probe pairs for each of

approximately 5000 full-length human genes. B. Probe design. To control for background and cross-hybridization, each perfect match probe is partnered with a probe of the same sequence except containing a central mismatch. Probes are usually 25mers, and are generally chosen to interrogate the 3' regions of eukaryotic transcripts to mitigate the consequences of partially degraded mRNA.

Please insert the following paragraph as paragraph 0018 after paragraph 0017:

Figure 11. Resequencing array for sequence variation detection. A. Each base of a given reference sequence is represented by four probes, usually 20mers, that are identical to each other with the exception of a single centrally located substitution (bold). Shown are probe sets targeted to two adjacent positions of the reference sequence. B. The target sequence is determined by hybridization intensities, with the probe complementary to the target providing the strongest signal.

Please insert the following paragraph as paragraph 0019 after paragraph 0018:

Figure 12. HuSNP array design. A. A known biallelic polymorphism at position 0 is interrogated by a block of four or five probe sets (five in this example). Each probe set consists of four probes, a perfect match and a mismatch to allele A, and a perfect match and a mismatch to allele B. One probe set in a block is centered directly over the polymorphism ("0"), and others are centered upstream (-4, -1) and downstream (+1, +4). B. The sequences of the probe set centered over the polymorphism is shown. C. Sample

images of blocks showing homozygous A, heterozygous A/B, or homozygous B at the same SNP site.

Please replace paragraph 0020 with the following amended paragraph:

Figure 13[[0]]. Schematic of the single-base extension assay applied to Tag probe arrays. Regions containing known SNP sites (A or G in this example) are first amplified by PCR. The PCR product serves as the template for an extension reaction from a chimeric primer consisting of a 5' tag sequence and a 3' sequence that abuts the polymorphic site. The two dideoxy-NTPs that could be incorporated are labeled with different fluorophores; in this example ddUTP is incorporated in the case of the A allele, and ddCTP for the G allele. Multiple SBE reactions can be done in a single tube. The tag sequence, unique for each SNP, directs the extension products to a particular address on the Tag probe array. The proportion of a fluorophor at an address reflects the abundance of the corresponding allele in the original DNA.

Please replace paragraph 0021 with the following amended paragraph:

Fig 14[[1]]. Using Tag probe arrays to screen protein activity. To a protein-encoding mRNA a 5' tag sequence and a 3' ribosome-blocking sequence are attached (A). In a pool of such molecules, such as a randomly mutated gene library, each mRNA is paired with a unique tag and all have the same 3' sequence. Following in-vitro translation either on a microarray or in a test tube, the nascent protein remains attached to

the mRNA (B). During hybridization the tag directs each mRNA-protein to a particular address on the Tag probe array (C), where all the proteins are screened simultaneously for activity (D). Appropriate detection methods identify proteins of interest (E, black and/or shaded blocks). Finally, the corresponding genes can be captured by PCR of the mRNA pool using a universal reverse primer and each identified Tag sequence as a forward primer.

Please replace paragraph 0022 with the following amended paragraph:

Figure 15[[2]]. PCR based method for attaching a tag sequence to a RNA. A gene sequence is hybridized with a forward primer which contains a T7 promoter, a tag sequence and Gene seq which is complementary with the gene sequence (A). A PCR results in a double stranded DNA that contains the gene sequence, the tag sequence and T7 promoter (B). An in vitro transcription reaction can be used to generate RNA that contains the coding region and the tag (C). The RNA can be used in vitro translation (D). The reverse primer for the PCR (A) contains both sequences for hybridizing with the gene sequence and a ribosome block sequence (Rblock). This block sequence can facilitate the retention of ribosome with the tagged RNA (D).

Please replace paragraph 0049 with the following amended paragraph:

Currently, the most popular application for oligonucleotide microarrays is in monitoring cellular gene expression. Standard GeneChip® arrays are encoded with

public sequence information, but custom arrays are also designed from proprietary sequences. Figure 10 depicts how a [[A]] gene expression array interrogates each transcript at multiple positions. This feature provides more accurate and reliable quantitative information relative to arrays which use a single probe, such as a cDNA clone or PCR product, for each transcript. Two probes are used at each targeted position of the transcript, one complementary (perfect match probe), and one with a single base mismatch at the central position (mismatch probe). The mismatch probe is used to estimate and correct for both background and signal due to non-specific hybridization. The number of transcripts evaluated per probe array depends upon chip size, the individual probe feature size, and the number of probes dedicated to each transcript. A standard 1.28 X 1.28 cm probe array, with individual 20 X 20 μ m features, and 16 probe pairs per probe set, can interrogate approximately 12,000 transcripts. This number is steadily increasing as manufacturing improvements shrink the feature size, and as improved sequence information and probe selection rules allow reductions in the number of probes needed for each transcript.

Please replace paragraph 0052 with the following amended paragraph:

Oligonucleotide arrays are currently used primarily for two types of genotyping analysis. *Arrays for mutation or variant detection* (Figure 11) are used to screen sets of contiguous sequence for single-nucleotide differences. Given a reference sequence, the basic design of genotyping arrays is quite simple: four probes, varying only in the central position and each containing the reference sequence at all other positions, are made to

interrogate each nucleotide of the reference sequence. The target sequence hybridizes most strongly to its perfect complement on the array, which in most cases will be the probe corresponding to the reference sequence, but in the case of a nucleotide substitution, this will be one of the other three probes. The other main type of genotyping performed with oligonucleotide arrays is *SNP analysis*, that is, the genotyping of biallelic single-nucleotide polymorphisms. Because SNPs are the most common source of variation between individuals, they serve not only as landmarks to create dense genome maps but also as markers for linkage and loss of heterozygosity studies. Large numbers of publicly available SNPs – nearly one million to date – have been found using gel-based sequencing as well as mutation detection arrays.

Please replace paragraph 0053 with the following amended paragraph:

In addition to mutation detection arrays, at least two other types of oligonucleotide arrays can be used for SNP analysis. The “HuSNP” assay allows nearly 1500 SNP-containing regions of the human genome to be amplified in just 24 multiplex PCRs and then hybridized to a single HuSNP array. The SNPs cover all 22 autosomes and the X chromosome. The probe strategy for a SNP array is shown in (Fig.12). The probes for each SNP on the HuSNP array interrogate not only the two alleles of the SNP position, but also 3 or 4 positions flanking the SNP; the redundant data are of higher quality for the same reasons that the use of multiple probes improves gene expression monitoring array data.

Please replace paragraph 0054 with the following amended paragraph:

Although it is anticipated that the HuSNP assay will be appropriate for many applications, a more generic alternative is available in the form of the GenFlex™ array. For this array, two thousand 20mer "tag" probe sequences were selected on the basis of uniform hybridization properties and sequence specificity. The array includes 3 control probes for each tag (a complementary probe and single-base mismatch probes for both the tag and its complement). One way to use the GenFlex™ array for SNP analysis is illustrated in (Fig. 13[[0]]). In this example, a single-base extension reaction is used, in which a primer abutting the SNP is extended by one base in the presence of the two possible dideoxy-NTPs, each of which is labeled with a different fluorophor. Since each target-specific primer carries a different tag, the identity of each SNP is determined by hybridization of the single-base extension product to the corresponding tag probe in the GenFlex™ array. The flexibility of the GenFlex™ approach lies in the freedom to partner any primer with any tag, a feature which enables other applications as well.

Please replace paragraph 0059 with the following amended paragraph:

In one aspect of the invention, methods are provided for the use of microarrays for proteomics and other protein screening applications. For example, by attaching a different oligonucleotide sequence tag to each member of a group of proteins to be analyzed, hybridization would allow them to be arrayed in discrete locations on a chip for parallel screening. Proteins of interest would be identified by their position on the array. In one exemplary approach (Fig. 14[[1]]), the tag is attached to the protein genetically by linking the tag to the mRNA and then translating the protein in such a manner that the

protein remains associated with the mRNA, as is done in ribosome display to create and capture high affinity antibodies (Hanes J, Jermutus L, Pluckthun A (2000) *Methods Enzymol* 328:404).

Please replace paragraph 0060 with the following amended paragraph:

A unique tag sequence can be attached to each target (mRNA, cDNA, gene, DNA fragment) in several ways. One method, depicted in Figure 15[[2]], incorporates a tag in a target-specific PCR primer, in this example, the forward primer. The forward primer for each target is assigned a different tag. Tagging n targets thus requires n different forward primers; the reverse primers can be either target-specific as in the example, or common to all targets if the targets have common ends, for example polyA tracts or adaptor attachments. Each target can be tagged in a separate PCR, or multiple reactions can be done in the same vessel, i.e., multiplex. As the figure depicts, additional features for transcribing and translating the target can be incorporated into the PCR primers.

Please replace paragraph 0063 with the following amended paragraph:

In some cases it is not necessary for each different target to have a unique tag. For example, in screening a library of protein variants, as depicted in Figure 14[[1]], in some cases it is acceptable for multiple variants to travel to the same address on the array. During screening the output signal from such an address is less pure than from an address with just one variant, and potential high signal can be diluted, but this drawback can be an acceptable trade-off depending on other conditions and throughput requirements.

Subsequent amplification of the targets on such an address can capture undesired variants, but an additional subsequent retagging and rescreening of all the captured variants makes it unlikely that the same unwanted variant is captured again. In other words, in some cases it can be more efficient to retag and rescreen than to require unique tags for each target.

In the Drawings:

Please reinstate Figures 10 through 12.

Figures 10 through 12 have been amended to add sequence identifiers (SEQ ID NO:) in order to comply with 37 C.F.R. 1.821.

Figures 13, 14 and 15, which had been renumbered as Figures 10, 11 and 12 respectively in a Preliminary Amendment, have been amended to restore the original numbering.

A complete new set of drawings is enclosed attached herewith.

Attachment: Replacement Sheets